



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Quantification, prediction and the online impact of sentence truth-value

Citation for published version:

Nieuwland, M 2016, 'Quantification, prediction and the online impact of sentence truth-value: Evidence from event-related potentials', *Journal of Experimental Psychology: Learning, Memory, and Cognition*, pp. 316-334. <https://doi.org/10.1037/xlm0000173>

Digital Object Identifier (DOI):

[10.1037/xlm0000173](https://doi.org/10.1037/xlm0000173)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Journal of Experimental Psychology: Learning, Memory, and Cognition

Publisher Rights Statement:

© Nieuwland, M. (2015). Quantification, prediction and the online impact of sentence truth-value: Evidence from event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*./ This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Quantification, prediction and the online impact of sentence truth-value:
Evidence from event-related potentials

Mante S. Nieuwland

Department of Psychology, School of Philosophy, Psychology and Language Sciences,
University of Edinburgh, Edinburgh, United Kingdom

Correspondence:

Mante S. Nieuwland
Department of Psychology
School of Philosophy, Psychology and Language Sciences
University of Edinburgh
7 George Square
Edinburgh EH8 9JZ
Scotland, United Kingdom
Email: m.nieuwland@ed.ac.uk
Telephone: +44 (0) 131 650 8387

ABSTRACT

Do negative quantifiers like ‘few’ reduce people’s ability to rapidly evaluate incoming language with respect to world knowledge? Previous research has addressed this question by examining whether online measures of quantifier comprehension match the ‘final’ interpretation reflected in verification judgments. However, these studies confounded quantifier valence with its impact on the unfolding expectations for upcoming words, yielding mixed results. In the current ERP study, participants read negative and positive quantifier sentences matched on cloze probability and on truth-value (e.g., “Most/Few gardeners plant their flowers during the spring/winter for best results”). Regardless of whether participants explicitly verified the sentences or not, true-positive quantifier sentences elicited reduced N400s compared to false-positive quantifier sentences, reflecting the facilitated semantic retrieval of words that render a sentence true. No such facilitation was seen in negative quantifier sentences. However, mixed-effects model analyses (with cloze value and truth-value as continuous predictors) revealed that decreasing cloze values were associated with an interaction pattern between truth-value and quantifier, whereas increasing cloze values were associated with more similar truth-value effects regardless of quantifier. Quantifier sentences are thus understood neither always in two sequential stages, nor always in a partial-incremental fashion, nor always in a maximally incremental fashion. Instead, and in accordance with prediction-based views of sentence comprehension, quantifier sentence comprehension depends on incorporation of quantifier meaning into an online, knowledge-based prediction for upcoming words. Fully incremental quantifier interpretation occurs when quantifiers are incorporated into sufficiently strong online predictions for upcoming words.

INTRODUCTION

Expression of quantity is a crucial aspect of people's ability to refer to things in the world in everyday language. Quantifiers are particularly important to differentiate between entities that belong to a larger set, and to communicate how representative a certain property is for the complete set of entities. The relationship between quantifier expressions and sentence truth-conditions has played a crucial role in formal semantics (e.g., Barwise & Cooper, 1981; Partee, 1996), in pragmatics (e.g., Geurts, 2010; Katsos & Cummins, 2010), and in the cognitive psychology of reasoning (e.g., Johnson-Laird, 1977). However, the impact of positive and negative quantifier expressions (e.g., 'most'/'few', respectively) on online language comprehension is still a matter of debate. This debate has centred on whether or not quantifier expressions are understood fully incrementally, that is, whether or not the initial comprehension of quantifier phrases (as reflected in online measures such as ERP or eye-tracking) matches their 'final' interpretation (as reflected in offline judgments). Some findings suggest that quantifier meaning does not impact the initial stages of semantic comprehension at all (e.g., Kounios & Holcomb, 1992), whereas others have reported that quantifier interpretation is neither fully immediate nor fully delayed (e.g., Urbach & Kutas, 2010). However, as I discuss in detail below, previous work has confounded quantifier valence with its impact on the unfolding expectations for upcoming words. The current study examines whether the incremental nature of quantifier comprehension depends on their incorporation into an online linguistic prediction about upcoming words.

It is widely recognized that incremental language comprehension is limited as listeners and readers do not always immediately interpret linguistic information to the fullest degree possible (e.g., Garrod & Sanford, 1999). This may also be the case for quantifiers given that the final interpretation depends on computation of quantifier scope (Kurtzman & MacDonald, 1993), the domain to which the quantifier needs to be applied. With each new bit of the unfolding input, the *prima facie* plausibility of the expression can change. The phrase "Few mayors see" seems implausible, while "Few mayors see ghosts" seems more plausible, and vice versa for the quantifier 'most'. To avoid misinterpretation, people generally hold off computing phrasal plausibility at 'see' when they anticipate a wider quantifier scope.

However, even when scope is unambiguous, the final interpretation might not be available immediately. Evidence for such a delay in quantifier comprehension has come from the N400 event-related potential, a negative voltage deflection whose amplitude peaks approximately 400 ms post-stimulus and indexes the extent to which retrieval of semantic memory associated with a word is facilitated by the context (Kutas & Hillyard, 1980). Kounios and Holcomb (1992) reported that in quantified sentences such as "All/No rubies are gems/spruces", 'gems' elicits a smaller N400 than 'spruces' irrespective of sentence truth-value ratings. Such findings suggest that 'gems', due to semantic category priming from 'rubies', is facilitated irrespective of the quantifier (see also Fischler, Bloom, Childers, Roucos, & Perry, 1983; Ferguson, Sanford & Leuthold, 2008). Such findings are also compatible with sentence verification models in which an affirmative proposition ("rubies are gems") is computed before applying the negative quantifier meaning (i.e., sequential interpretation; Carpenter & Just, 1975).

Urbach and Kutas (2010), however, reported different results for quantifier phrases with sentential objects that were typical or atypical for the preceding verb (e.g., "Few/Most mayors see citizens/ghosts on a regular basis"). Post-sentence plausibility ratings showed that participants found typical objects more plausible than atypical objects following positive quantifiers, but less plausible following negative quantifiers. Atypical objects elicited the same N400s following positive and negative quantifiers. However, while typical objects elicited smaller N400s overall, N400s were more reduced following positive quantifiers compared to negative quantifiers. Thus, quantifier meaning impacted how the objects were

processed as soon as they appeared, but the online effects of quantifier meaning differed qualitatively from their effects on final interpretation. Urbach and Kutas argued that quantifier comprehension was incremental yet partial, meaning that initial semantic processes were influenced by quantifier meaning yet not to the extent that they delivered an interpretation that was compatible final evaluation of plausibility.

The few available studies on quantifier comprehension thus leave open why comprehension sometimes proceeds incrementally yet partially, and whether comprehension can also proceed fully incrementally (i.e., when patterns from online and offline measures align). Partially incremental comprehension of quantifiers may be an epiphenomenon of the context in which quantifiers are used. As Urbach and Kutas note, sentences like “Few mayors see ghosts” lack a supportive pragmatic context, as usually available in natural language settings. Atypical objects are inherently less semantically related to the context words, and are less predictable from the sentence context. Hence, while the negative quantifier ‘few’ may decrease expectations for ‘citizens’, the expectation for ‘ghost’ may not increase substantially. The things in the world that only few mayors see are practically unlimited (or at least substantially more numerous than things that most mayors see). The phrases “Few mayors see” and “Most mayors see” may both lead readers to expect a mayor-related concept. The object ‘ghosts’ will be equally unexpected following negative and positive quantifiers. The results from Urbach and Kutas may reflect the extent to which the objects were expected or unexpected given the context¹, not an inherent asymmetry in how positive and negative quantifiers are understood online (see Nieuwland & Kuperberg, 2008, for related discussion on negation comprehension). In explaining their findings, previous studies have overlooked the role of predictability of the critical word, as is typically assessed with cloze probability scores. Higher cloze values suggest increased likelihood of online predictive processing, which in turn facilitates incremental processing (e.g., Altmann & Mirković, 2009; Staub & Clifton, 2006). Comprehension of positive and negative quantifier sentences may become increasingly similar when those sentences elicit relatively strong linguistic predictions about upcoming words.

The current study

The current study disentangled the impact of quantifier valence and of predictability by comparing ERPs to positive and negative quantifier sentences that are pre-tested on cloze-value, truth-value, and on semantic relatedness of the critical words to the sentence context (Table 1). Critical words came in pairs so that one word rendered the negative quantifier-sentence true and the positive quantifier-sentence false, and the reverse for the other word. Sentence materials were used with varied cloze values. A linear mixed-effects model analysis with continuous predictors for single-trial N400 data was performed to predict the interaction pattern between truth-value and quantifier based on item cloze value. Predictions centred on finding smaller N400s for words that render sentences true compared to words that render sentences false (i.e., sentence truth-value N400 effects). Such effects arise from world knowledge-based predictions about upcoming words (Nieuwland & Kuperberg, 2008; Nieuwland & Martin, 2012). Here, I examine sentence truth-value N400 effects when participants are explicitly verifying the sentences and when they are not. Evidence for partial incrementality can be considered stronger if obtained when participants are engaged in

¹ Their conclusions were also limited by the fact that participants judged sentences like “Few mayors see ghosts” to be neither plausible nor implausible (average ERPs were computed over sentences that were regarded as plausible or implausible), and by the lack of direct correspondence between mid-sentence ERPs to post-sentence plausibility judgments, where the perceived quantifier scope may have differed.

explicit sentence verification. Moreover, effects that occur regardless of whether participants are engaged in explicit sentence verification cannot solely be ascribed to strategic task-effects.

Given that the N400 indexes the extent to which a word receives semantic facilitation from context-based predictions, and does not directly index the online computation of truth-value, several hypotheses can be formulated. If knowledge-based predictions impact comprehension in negative and positive quantifiers alike (see Nieuwland & Kuperberg, 2008), a similar N400 reduction for true compared to false sentences may be observed in positive and negative quantifier sentences. Alternatively, if negative quantifiers compromise the ability to generate online predictions for upcoming words, a smaller N400 reduction is observed in negative quantifier sentences than in positive quantifier sentences (Urbach & Kutas, 2010), or possibly no impact of quantifier type or truth-value on the N400 altogether (Kounios & Holcomb, 1992). Importantly, however, if incremental comprehension of quantifier sentences depends on people's usage of quantifiers to generate linguistic predictions, then effects of truth-value in negative and positive quantifier sentences are more similar in high-cloze sentences. These hypotheses are tested with a mixed-effects model analysis that includes cloze value and truth-value as continuous predictors and quantifier (positive/negative) as a dichotomous predictor for the dependent variable N400 amplitude.

METHODS

Participants

Written informed consent was obtained from 60 right-handed Edinburgh University students (21 males) between 19 and 35 years old (30 participants per experiment). All were native English speakers, and none had neurological or psychiatric disorders or participated in the pre-tests.

Materials

Two hundred sentence quadruplets were constructed that started with quantifier expressions similar to the ones used by Urbach and Kutas (2010). Critical words (CWs) per quadruplet came in pairs, with one word rendering a negative quantifier sentence true and a positive quantifier sentence false, and the reverse for the other word. The sentences covered different world-knowledge topics. Further item selection was based on two pre-tests. First, from one of two counterbalanced lists, 28 participants completed either the positive or the negative version of each sentence truncated before the CW, using the first sensible word coming to mind. Cloze value was computed as the percentage of participants who used the intended CW. Second, from one of four counterbalanced lists, 40 different participants evaluated one condition of each quadruplet on truth-value (1 = False, 5 = True).

Quadruplets were excluded when containing a true sentence with an average rating below 3 or a false sentence with an average rating over 3. Further selection was done so that the number of items was a multiple of four, and to minimize differences between conditions on the matching variables. In the final set of 124 items (available in the Appendix), critical word-pairs differed slightly in lexical frequency ($t(123) = 1.7, p = .08$), but were matched on lexical co-occurrence with words in the sentence (LSA-SSV, <http://lsa.colorado.edu>; $t(123) = 1.3, p = .19$). A 2(truth-value: true, false) by 2(quantifier: positive, negative) 2-way interaction analysis of the cloze values revealed that true sentences had higher cloze values than false sentences ($F(1,123) = 226.4, p < .001$), and that negative sentences had higher cloze values than positive sentences ($F(1,123) = 5.9, p = .017$), but no robust interaction between truth-value and quantifier type ($F(1,123) = 0.23, p = .88$). A similar analysis for the truth-value ratings showed that true/false ratings differed between the positive and negative quantifier sentences ($F(1,123) = 17.1, p < .001$), due to the fact that true-positive sentences were rated as slightly higher than true-negative sentences ($t(123) = 2.1, p = .04$), whereas false-positive sentences were rated as slightly lower than false-negative sentences ($t(123) = 2.9, p = .004$). When considering these numerically very small but statistically robust differences in cloze values and truth-value, however, it is important to keep in mind that the analysis here uses truth-value and cloze value as continuous predictors, rendering condition averages less relevant.

In the ERP experiment, CWs were followed by three additional words. Four counterbalanced lists were created so that each sentence appeared in only one condition per list, but in all conditions equally often across lists. Within each list, items were pseudorandomly mixed with 216 filler sentences (126 true and 90 false, see Nieuwland, 2015).

Procedure

Participants silently read sentences, presented word-by-word, while minimizing movement. Word duration was 300 ms, and the inter-word-interval was 200 ms. Sentence-final words were presented for 600 ms and followed by a blank screen for 1800 ms.

In the verification-instruction, a response display followed showing the response options 1-2-3-4-5 centred on the screen and “Strongly disagree” and “Strongly agree” below the 1 and 5, respectively. Participants were instructed to respond as accurately as possible using the keyboard without time limit. After the response, a fixation mark appeared indicating that participants could self-pace on to the next sentence by pressing the space bar.

In the no-verification instruction, the post-sentence blank screen was followed either by a fixation mark or by a yes/no world-knowledge question to which participants answered by button-press (followed by the fixation mark). These questions (82 in total, 41 requiring a ‘yes’ response) were only included to keep participants alert, and they were independent of sentence condition and identical in each list.

Electroencephalogram recording and data processing

The electroencephalogram (EEG) was recorded at 64 EEG electrodes, along with two mastoid electrodes and four EOG electrodes. The EEG was re-referenced offline to the mastoid-average, filtered (0.05-30 Hz), segmented into epochs from -200 to + 1500 ms relative to CW onset, corrected for eye-movements using independent component analysis, baseline-corrected to -200 to 0 ms, and screened for artefacts (maximal/minimal amplitude at +100/-100 μ V). In the verification-condition, only trials with correct responses (strongly disagree or disagree for sentences pre-rated as being false, agree or strongly agree for sentences pre-rated as being true) were used for further analysis, which yielded an average of either 26 or 27 trials per condition. Analysis of the truth-value ratings obtained during the ERP experiment (see Table 1) showed that true sentences were rated higher than false sentences ($F(1,26) = 4906.2, p < .001$) and that positive sentences were rated as slightly lower than negative sentences ($F(1,26) = 5.4, p = .03$), but that the true-false rating differences were similar in positive and negative sentences ($F(1,26) = 0.1, p = .75$).

Participants were excluded from analysis if more than 1/3 of trials were rejected due to artefacts or condition-inconsistent responses (verification-instruction) or due to artefacts only (no-verification). This left 51 participants for the analysis (27 participants that did verification, 24 that did not; average number of trials, *verification*: true-positive, $M = 25.7, SD = 2.6$; false-positive, $M = 24.6, SD = 3.2$; true-negative, $M = 25.2, SD = 3.0$; false-negative, $M = 24.0, SD = 3.3$; *no-verification*, true-positive, $M = 27.8, SD = 3.4$; false-positive, $M = 28.5, SD = 3.8$; true-negative, $M = 28.3, SD = 3.4$; false-negative, $M = 27.8, SD = 4.4$). All subjects had at least 18 trials in each of the conditions.

Statistical analysis

Mixed-effect model analyses (Baayen, Davidson & Bates, 2008) were performed using the ‘lme4’ package (Bates, Maechler, Bolker & Walker, 2014) in the R software (R Core Team, 2014). The dependent variable was mean amplitude in the 300-450 ms (N400) time window at 22 posterior electrodes (CP1/3/5, P1/3/5/7, TP7, O1, PO3/7 and right equivalents), based on previously observed posterior N400 modulations (e.g., Nieuwland & Kuperberg, 2008; Nieuwland & Martin, 2012). In all analyses, resulting t-values of 2.00 and above are treated as significant, due to complexities in estimating the degrees of freedom associated with predictors (Baayen, 2008). A first linear mixed-effect model was constructed that included all 3-way interaction terms between verification-instruction, quantifier-type, and the two continuous predictors (pre-rated) truth-value and cloze value². Cloze value here

² Because the 4-way interaction (verification*clozevalue*quantifier*truthvalue) did not yield a significant effect, the model comparison between Model 1 and Model 2 was simplified by leaving out the 4-way interaction term.

Model 1: $N400 \sim (\text{verification} + \text{clozevalue} + \text{quantifier} + \text{truthvalue})^3 + (1 + \text{clozevalue} * \text{quantifier} * \text{truthvalue} \mid \text{subject}) + (1 + \text{quantifier} * \text{truthvalue} \mid \text{item})$.

Model 2: $N400 \sim (\text{verification} + \text{clozevalue} + \text{quantifier} + \text{truthvalue})^3 - \text{clozevalue} : \text{quantifier} : \text{truthvalue} + (1 + \text{clozevalue} * \text{quantifier} * \text{truthvalue} \mid \text{subject}) + (1 + \text{quantifier} * \text{truthvalue} \mid \text{item})$.

refers to the average cloze value of the true positive and true negative sentences, a procedure that was used to circumvent co-linearity between the cloze value and truth-value due to the selection of false sentences with cloze values near zero. As random effects, the model included intercepts for subjects and items, a by-subject random slope for the interaction between quantifier-type, truth-value and cloze value, and a by-item random slope for the interaction between quantifier-type and truth-value³. With a likelihood ratio test using an ANOVA, this model was compared to a second model that included the same predictors and 3-way interactions but not the 3-way interaction term between cloze value, quantifier type and truth-value.

RESULTS

As shown in Figure 1, critical words in all conditions elicited a positive P2 component followed by a negative N400 component, whether or not participants were performing an explicit verification task. The mixed-effects model comparison revealed that the full interaction had a significantly better fit than a model without the 3-way interaction term between cloze value, truth-value and quantifier type ($\chi^2_{(1)} = 4.0$; $p = .046$). In the former model, the only robust 3-way interaction was observed between cloze value, truth-value and quantifier type ($t = 2.1$). The nature of this interaction is illustrated in the lower graph in Figure 1, which contains scatterplots of the mean fitted values and regression lines associated with the mixed effects model analysis, created with ggplot2 (Wickham, 2009). For ease of exposition, voltage is plotted as a function of quantifier type and truth-value rating in the set with relatively high and low cloze values separately (based on a median split of the cloze value predictor). In items with higher cloze values, N400s become smaller (less negative) with increasing truth-value irrespective of quantifier type. A similar pattern occurred for positive quantifier sentences with relatively low cloze values, however, the reverse pattern occurred for negative quantifier sentences with low cloze values. The 3-way interaction pattern was replicated in a mixed-effects model analysis that tested a more traditional factorial design using dichotomous predictors (negative/positive quantifier, true/false, low/high cloze value). The results of this analysis and the accompanying pairwise follow-ups are available as Supplementary Materials.

DISCUSSION

This study examined ERPs elicited by words that rendered positive and negative quantifier sentences true or false. Positive and negative quantifier sentences elicited different effects in the complete set of items. Critical words in true-positive quantifier sentences elicited smaller N400s than false-positive quantifier sentences, reflecting the facilitated semantic retrieval for words that render a sentence true, regardless of whether people explicitly verified the sentences or not (e.g., Nieuwland, 2013; for behavioral findings, see Singer, 2013). No such effect was observed in negative quantifier sentences. Critically, the observed effects depended on cloze value, such that decreasing cloze values were associated with an interaction pattern between truth-value and quantifier, whereas increasing cloze values were associated with more similar truth-value effects regardless of quantifier.

The current findings have several theoretical implications. First, the results are inconsistent with older accounts of quantifier comprehension. Kounios and Holcomb (1992)

³ Although an attempt was made to follow a fully maximal random effects structure, following Barr, Levy, Scheepers and Tily (2013), a model that included a by-item random slope for the three-way interaction between verification-instruction and truth-value and quantifier did not converge.

argued that, when people read quantifier sentences, they initially access semantic memory via non-propositional processes, and compute sentence truth-value through later, decision-based processes (see also Fischler et al., 1983). This prediction also followed from traditional sentence verification models, in which comprehension of negation involves the initial computation of an affirmative proposition followed by application of the negation term (Carpenter & Just, 1975). Both these accounts predict that the N400, an index of semantic retrieval processes (Kutas & Hillyard, 1980), is insensitive to quantifier valence, and is solely a function of the lower-level semantic relationships between critical words and context words. The current results show clearly that this explanation does not adequately describe how people comprehend sentences with different contextual constraints.

Second, the current results also shed further light on a recent proposal that quantifier phrases are understood immediately yet not completely incrementally. Urbach and Kutas (2010) reported that for sentences like “Few/Most mayors see ghosts/citizens”, semantic retrieval for typical objects like ‘citizens’ is facilitated (i.e., eliciting smaller N400s) following positive quantifiers compared to negative quantifiers, whereas atypical objects like ‘ghosts’ are not facilitated following negative quantifiers compared to positive quantifiers. While the plausibility ratings in that study suggested that the quantifier expressions were interpreted fully, i.e., consistent with theoretical accounts of their meaning, the N400 patterns at the critical words suggested that the meanings of quantifiers were initially registered were not fully incorporated into the evolving representation of the sentential context. The misalignment between online and offline measurements led Urbach and Kutas to conclude that quantifier phrases are not understood fully incrementally. Importantly, the current results show that this explanation too does not adequately describe how people comprehend sentences with different contextual constraints. Misalignment between online and offline measurements disappeared with increasing cloze values. The current results therefore suggest that fully incremental quantifier interpretation can be attained when quantifiers are successfully incorporated into predictions for upcoming words based on real-world knowledge.

The interaction pattern between quantifier type and truth-value as observed for lower cloze sentences is reminiscent of previous findings on quantifier comprehension (Kounios & Holcomb, 1992; Urbach & Kutas, 2010). Nieuwland and Kuperberg (2008) also reported a similar pattern for implausible negation (e.g., “A bulletproof vest is not dangerous”). However, in light of the current cloze probability and truth-value ratings, the current results seem not to arise from implausible use of negative quantifiers. In fact, the current set as a whole is more similar to the plausible negation sentences from Nieuwland and Kuperberg in terms of offline judgements, which elicited the same N400 sentence truth-value effects as affirmative sentences. So, the question arises why negative quantifiers would pose a particular challenge to incremental comprehension in plausible but low-cloze sentences.

One potentially relevant difference between the current sentences and those from Nieuwland and Kuperberg, beside the difference between negation and negative quantifiers, is the sentence position of the negative operator (sentence initial in the current study, right before the critical word in the Nieuwland and Kuperberg materials). Incremental incorporation of a negative term into the sentence representation may be particularly difficult when it occurs early in the sentence, when the quantifier scope is still completely undetermined. In contrast, when a negation term appears mid-sentence, the property that needs to be negated may already be sufficiently salient or presupposed by the preceding sentence material. Availability of the to-be-negated information is one of the major determinants of ease with negation comprehension (e.g., Kaup, Zwaan & Lüdtke, 2007;

Nordmeyer & Frank, 2014; Wason, 1965; Tian, Breheny, & Ferguson, 2010). In addition, it may be more difficult for people to predict the end of the quantifier scope than the end of a negation scope during comprehension. Such differences would be exacerbated by the word-by-word presentation procedure, which does not give the same information to predict scope end as would be available during natural reading or listening.

Prediction-guided full and incremental quantifier interpretation

Quantifier sentences seem to be understood neither always in two sequential stages, nor always in a partial-incremental fashion, nor always in a maximally incremental fashion (i.e., delivering the fullest interpretation of a sentence fragment at each moment of the fragment's unfolding; Altmann & Mirković, 2009). Instead, whether comprehension of quantifier sentences appears to occur in two stages, partial-incrementally or maximally incremental depends on the incorporation of quantifier meaning into an online prediction for an upcoming word. This could be a gradual phenomenon without a specific predictability tipping-point, although the required prediction strength remains to be established.

This study shows a clear impact of whether or not a quantified sentence allows readers to predict upcoming words, but it used a range of different sentence types and therefore does not directly address the question of what makes a quantified sentence more or less predictable. Various sources of information are likely to contribute, based on well-established findings in the literature on negation comprehension and on predictability more generally. Clearly, there are general constraints that would apply to quantified sentences and non-quantified sentences alike. Predictions are partly driven by real-world factual knowledge as well as the pragmatic presumption of relevance (e.g., Grice, 1975; Sperber & Wilson, 1986). By default, people do not expect speakers to utter a sentence that they do not mean or to utter a sentence that has no obvious relevance to the on-going discourse. This is what makes an isolated sentence such as “Few mayors see ghosts” or “No rubies are spruces” pragmatically infelicitous, but possibly very meaningful when placed in an appropriate context. In fact, very recent results from Urbach, Delong & Kutas (2015) suggest that a supportive discourse context can but does not always lead to incremental quantifier comprehension⁴. Another contribution comes from sentence constraints that limit the number of grammatically correct continuations that satisfy the abovementioned pragmatic criteria. For example, the sentence “Most/Few gardeners plant their flowers during the” could lead comprehenders to predict a temporal noun phrase (e.g., ‘weekend’, or a season). However, quantifiers may also generate more specific contributions to predictability by making a relevant contrast set available (e.g., Filik, Leuthold, Moxey, & Sanford, 2011; Paterson, Filik & Moxey, 2009). This might be particularly important for negative quantifier sentences, given the contextual use of negation in everyday language (Horn, 1984). For example, while “Few bananas are” will probably not elicit specific predictions, the phrase “Few green bananas are” may activate relevant real-world knowledge about the association between banana colour and its ripeness. If the relevant contrast set (ripe-unripe) becomes available during comprehension, an online prediction may be generated that facilitates the upcoming word ‘ripe’. Even an isolated negative quantifier sentence could then achieve a level of

⁴ Participants in one experiment read isolated sentences such as “Few/Most children prefer vegetables/sweets”, whereas participants in a second experiment read them after a supportive context sentence (e.g., “Alex was an unusual toddler”). Only with the context was a fully incremental pattern observed, although no between-experiment analysis was performed to directly test the effect of context. Surprisingly, the incremental pattern was not replicated in another experiment wherein participant evaluated plausibility of the sentence plus context.

required presupposition that mimics the contextualized use of negative quantifiers in regular language use (e.g., Nieuwland & Kuperberg, 2008; Tian et al., 2010).

At least some amount of prediction appears to be required for fully incremental comprehension of quantifiers. This fits with recent views that prediction plays an important role in language (Altmann & Mirković, 2009; Elman, 1990; Levy, 2008), and that incremental comprehension is an emergent property of a language system that tries to continuously predict upcoming information at various levels of representation (e.g., phonology, semantic meaning, event-knowledge). The current study thus shows that quantity information can be incorporated into such predictions of semantic information and their effects on online processing. As assumed by psycholinguistic studies on prediction, this claim relies on cloze probability as an independent measure of critical word predictability. This assumed link between an offline production measure and online semantic processing is not without its problems, however. Cloze values are obtained in absence of time constraints (although see Staub, Grant, Astheimer & Cohen, 2015), whereas online comprehension, especially in the standard serial visual presentation procedure, proceeds under time constraints. This means that there might be situations in which high cloze does not necessarily lead to online prediction-associated N400 effects (e.g., when words are presented at an uncomfortably rapid pace), and situations in which low-cloze negative sentences do lead to online predictions (e.g., when words are presented at a very slow pace or when participants are in control of the reading rate such as in natural sentence reading). Alternatively, the rich prosodic information available during auditory comprehension may also aid the online generation of predictions. Slower and/or more naturalistic presentation procedures may have a bigger impact on how people understand negative quantifiers than how they understand positive quantifiers.

CONCLUSION

Sentence truth-conditions have traditionally played an essential role in philosophical and linguistic theories of meaning. In psycholinguistics, however, people's ability to establish sentence truth-value has long been considered as reasoning about previously understood sentences (e.g., Hasson, Simmons & Todorov, 2005). This view has been reinforced by the often-reported insensitivity of early measures of semantic processing to sentence truth-value in negated sentences (e.g., Fischler et al., 1983; Kounios & Holcomb, 1992). Critically, the current study suggests that previously reported dissociation between online and offline measures of quantifier comprehension disappears when quantifiers are incorporated into an online linguistic prediction for specific words. Negative quantifiers like 'few' do not principally reduce people's ability to rapidly map incoming language onto what they hold to be true, as evidenced by sentence truth-value N400 effects in positive and negative quantifier sentences alike. These N400 effects reflect the facilitated semantic retrieval for words that are consistent with knowledge-driven online predictions. These predictions are what propel fully incremental quantifier comprehension.

REFERENCES

- Altmann, G., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, 33(4), 583-609.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390-412.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255-278.
- Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4, 159-219.
- Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. (R package version 1.1-7). [Computer software].
- Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: a psycholinguistic processing model of verification. *Psychological Review*, 82(1), 45.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14(2), 179-211.
- Ferguson, H. J., Sanford, A. J., & Leuthold, H. (2008). Eye-movements and ERPs reveal the time course of processing negation and remitting counterfactual worlds. *Brain Research*, 1236, 113-125.
- Filik, R., Leuthold, H., Moxey, L. M., & Sanford, A. J. (2011). Anaphoric reference to quantified antecedents: An event-related brain potential study. *Neuropsychologia*, 49(13), 3786-3794.
- Fischler, I., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry, N. W. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, 20(4), 400-409.
- Garrod, S., & Sanford, A. (1999). Incrementality in discourse understanding. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading*. (pp. 3-27). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Geurts, B. (2010). *Quantity implicatures*. Cambridge University Press.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41-58). New York: Academic Press.
- Hasson, U., Simmons, J. P., & Todorov, A. (2005). Believe It or Not On the Possibility of Suspending Belief. *Psychological Science*, 16(7), 566-571.
- Horn, L. (1989). *A natural history of negation*. Chicago: University of Chicago Press.
- Johnson-Laird, P. N. (1977). *Reasoning with quantifiers*. Thinking. Cambridge University Press, Cambridge.
- Katsos, N., & Cummins, C. (2010). Pragmatics: from theory to experiment and back again. *Language and Linguistics Compass*, 4(5), 282-295.
- Kaup, B., Zwaan, R.A., & Lüdtke, J. (2007). The experiential view of language comprehension: How is negated text information represented? In F. Schmalhofer & C.A. Perfetti (Eds.), *Higher level language processes in the brain: Inference and comprehension processes* (pp. 255-288). Mahwah, NJ: Erlbaum.
- Kounios, J., & Holcomb, P. J. (1992). Structure and process in semantic memory: evidence from event-related brain potentials and reaction times. *Journal of Experimental Psychology: General*, 121(4), 459-479.
- Kurtzman, H. S., & MacDonald, M. C. (1993). Resolution of quantifier scope ambiguities. *Cognition*, 48(3), 243-279.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, 207(4427), 203-205.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106, 1126-1177.

- Nieuwland, M.S. (2013). "If a lion could speak ...": Online sensitivity to propositional truth-value of unrealistic counterfactual sentences. *Journal of Memory and Language*, 68(1), 54-67.
- Nieuwland, M.S. (2015). The truth before and after: Brain potentials reveal automatic activation of event-knowledge during sentence comprehension. Manuscript accepted pending revisions.
- Nieuwland, M. S., Ditman, T., & Kuperberg, G. R. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language*, 63(3), 324-346.
- Nieuwland, M. S., & Kuperberg, G. R. (2008). When the truth is not too hard to handle an event-related potential study on the pragmatics of negation. *Psychological Science*, 19(12), 1213-1218.
- Nieuwland, M. S., & Martin, A. E. (2012). If the real world were irrelevant, so to speak: The role of propositional truth-value in counterfactual sentence comprehension. *Cognition*, 122(1), 102-109.
- Nordmeyer, A. E., & Frank, M. C. (2014). The role of context in young children's comprehension of negation. *Journal of Memory and Language*, 77, 25-39.
- Paterson, K. B., Filik, R., & Moxey, L. M. (2009). Quantifiers and discourse processing. *Language and Linguistics Compass*, 3(6), 1390-1402.
- Partee, B. H. (1996). The Development of Formal Semantics in Linguistic Theory. In *Handbook of Contemporary Semantic Theory*, ed. S. Lappin, 11-38. Oxford: Blackwell.
- R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Singer, M. (2013). Validation in reading comprehension. *Current Directions in Psychological Science*, 22(5), 361-366.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Basil Blackwell.
- Staub, A., & Clifton Jr, C. (2006). Syntactic prediction in language comprehension: evidence from either... or. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(2), 425-436.
- Staub, A., Grant, M., Astheimer, L., & Cohen, A. (2015). The influence of cloze probability and item constraint on cloze task response time. *Journal of Memory and Language*, 82, 1-17.
- Tian, Y., Breheny, R., & Ferguson, H. J. (2010). Why we simulate negated information: A dynamic pragmatic account. *The Quarterly Journal of Experimental Psychology*, 63(12), 2305-2312.
- Urbach, T. P., DeLong, K. A., & Kutas, M. (2015). Quantifiers are incrementally interpreted in context, more than less. *Journal of Memory and Language*, 83, 79-96.
- Urbach, T. P., & Kutas, M. (2010). Quantifiers more or less quantify on-line: ERP evidence for partial incremental interpretation. *Journal of Memory and Language*, 63(2), 158-179.
- Wason, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, 4(1), 7-11.
- Wickham, H. (2009). *Ggplot2: elegant graphics for data analysis*. Springer New York.

FIGURE LEGEND

Figure 1. Grand-average ERPs at electrode CP1 elicited by critical words in true and false sentences containing positive and negative quantifier expressions, across all participants and for the separate verification instructions (upper graphs). These ERP waveforms are high cut-off filtered at 5 Hz for presentation purposes, and negativity is plotted upwards. An example sentence in each condition is shown below the ERP waveforms, along with the scalp distribution of the overall sentence truth-value effect (false minus true) for positive and negative quantifier sentences. The lower graphs illustrate the nature of the interaction pattern between cloze value, pre-tested truth-value and quantifier type. For ease of exposition, these graphs show the mean fitted values from the mixed effects model results separately for sentences with high cloze values and low cloze values as scatterplots.

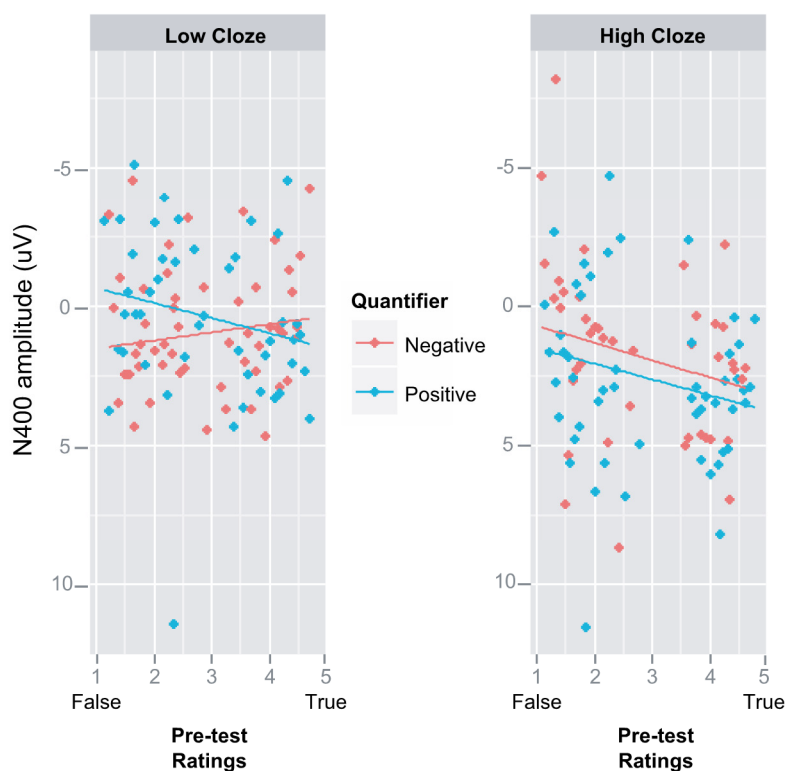
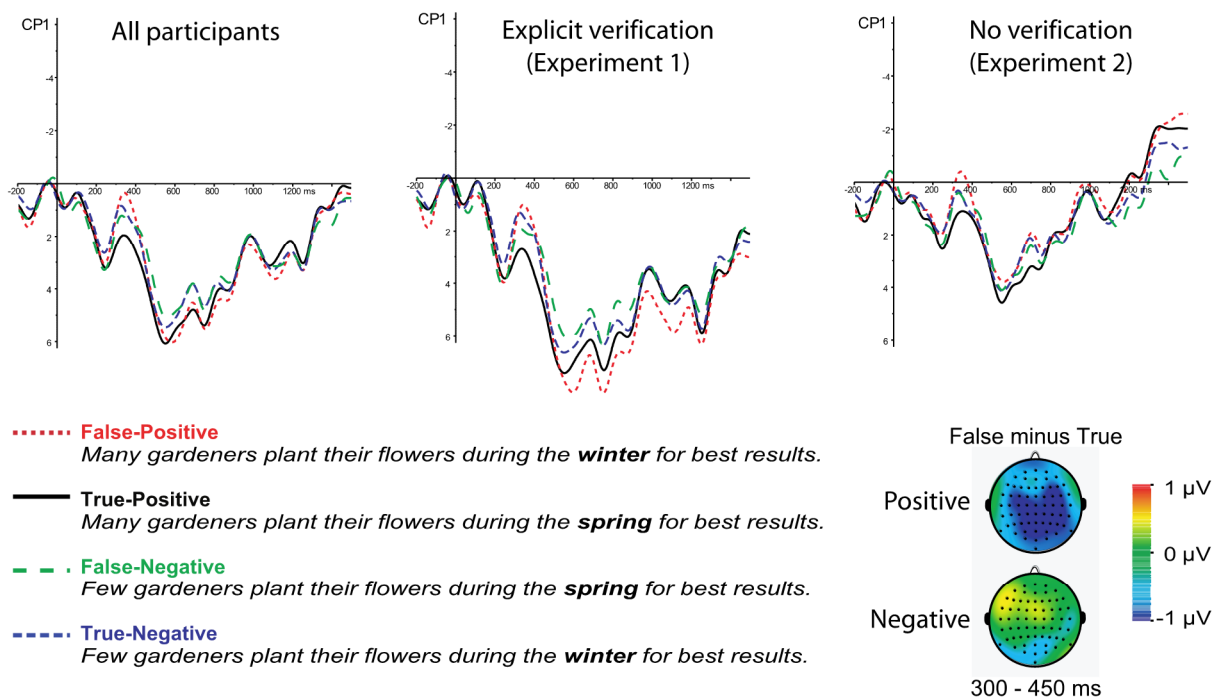
ACKNOWLEDGEMENTS

I thank Cassandra Addai, Keelin Murray, Chrysa Retsa, Aine Ito and Rachel King for help with materials and data collection. I also thank Andrea Eyleen Martin, Heather Ferguson and two anonymous reviewers for their comments on a previous draft of this manuscript, and Martin Corley for statistical advice. British Academy grant SG131266 funded this work.

Table 1. Example sentences with results from the cloze value pre-test, from the truth-value pre-test, and from the truth-value rating during the ERP experiment. Additional characteristics of the critical words are also given. The full set of materials is available as Appendix.

Condition	Example sentences	Cloze value (%)	Pre-rated truth-value	Truth-value rating in the experiment	Length in letters	Log frequency	Lexical co-occurrence
True-Negative	Few gardeners plant their flowers during the <u>winter</u> for best results. Hardly any retirees report that their health is starting to <u>improve</u> as they age.	34 (24)	4.05 (.32)	4.69 (.15)	5.7 (2.1)	1.63 (.73)	.18 (.07)
True-Positive	Many gardeners plant their flowers during the <u>spring</u> for best results. Lots of retirees report that their health is starting to <u>decline</u> as they age.	32 (26)	4.11 (.34)	4.65 (.15)	6.4 (2.0)	1.49 (.82)	.19 (.08)
False-Negative	Few gardeners plant their flowers during the <u>spring</u> for best results. Hardly any retirees report that their health is starting to <u>decline</u> as they age.	4 (8)	1.91 (.38)	1.28 (.15)	6.4 (2.0)	1.49 (.82)	.19 (.08)
False-Positive	Many gardeners plant their flowers during the <u>winter</u> for best results. Lots of retirees report that their health is starting to <u>improve</u> as they age.	2 (4)	1.79 (.39)	1.23 (.15)	5.7 (2.1)	1.63 (.73)	.18 (.07)

Note. Standard deviations are given in parentheses. Critical words are underlined for expository purposes. For truth-value ratings, 1 = False, 5 = True. Log frequency is based on the Celex corpus (celex.mpi.nl). Lexical co-occurrence is indexed with C-context semantic similarity values obtained with Latent Semantic Analysis (<http://lsa.colorado.edu>), reflecting the average co-occurrence of the critical word with each word in the sentence context.



Appendix. Sentences 1 to 124 correspond to the positive/negative true/false quantifier sentences. The negative quantifiers consisted predominantly of quantifiers that license negative polarity items such as ‘ever’ or ‘hardly’, with the exception of a small number of the items (e.g., ‘a minority of’, ‘a small number of’, 9 out of the final set of 124 sentences). Frequency count of used positive quantifier expressions: A large majority of (1), A large number of (7), A lot of (2), All (1), Almost all (16), Lots of (29), Many (11), Most (23), Mostly all (1), Nearly all (14), Practically all (14), The majority of (1), Very many (4). Frequency count of used negative quantifier expressions: A minority of (1), A small minority of (1), A small number of (7), A very few (2), Almost no (14), Few (36), Hardly any (15), No (1), Practically no (15), Rather few (14), Very few (18). The median split column lists each item based on the average cloze value of positive and negative quantifier sentences, which corresponds to the item split shown in in Figure 1 (HC = high cloze, LC = low cloze).

Item	Sentence conditions	Cloze Values (0-1)					Pre-rated truth-value			
		True-Positive	False-Positive	True-Negative	False-Negative	Median split	True-Positive	False-Positive	True-Negative	False-Negative
1	Practically all/Practically no postmen prefer delivering mail when the weather is good/bad during the day.	0.27	0	0.4	0	HC	4.5	1.15	4.38	1.85
2	Nearly all/Rather few doctors treat cancer with medicines that are strong/weak as a preference.	0.13	0	0.13	0	LC	4.08	1.69	3.58	1.62
3	Lots of/Hardly any recipes that are used for children's cooking classes are simple/difficult in the kitchen.	0.47	0	0.53	0	HC	4.5	1.54	4.15	1.85
4	Many/Few archaeologists find artefacts that are completely damaged/intact in the desert.	0.07	0.13	0.33	0	LC	3.69	1.69	4.23	2.58
5	Most/Few farmers grow crops that are harvested in the summer/winter in Great Britain.	0.4	0.07	0.87	0.07	HC	3.75	1.62	3.85	2.31
6	Practically all/Practically no lions are very happy living in the wild/zoo all their life.	0.73	0	0.47	0	HC	3.85	2	4	1.33
7	Nearly all/Rather few artists think that the economic crisis will cause their income to decrease/increase in this year.	0.6	0	0.47	0.27	HC	3.77	2.17	3.77	2.23
8	Lots of/Hardly any historians study articles that they believe to be fascinating/fake for their work.	0.07	0	0.2	0	LC	4.38	1.85	3.92	1.92
9	Lots of/Very few novels by aspiring authors become an instant flop/classic in this market.	0.07	0.07	0.07	0	LC	3.31	2.15	4.17	2.23
10	Most/Few cats during the winter like to sleep in spots that are warm/cold in the daytime.	0.8	0	0.73	0	HC	4.31	1.46	4.54	1.5
11	Almost all/Almost no caretakers clean floors at times when the place is quiet/busy after working hours.	0.2	0	0.53	0	HC	4.31	1.85	4	1.62
12	Many/Few lawyers in the UK think that reinstating the death penalty is negative/positive for modern society.	0.07	0	0.13	0	LC	4	1.62	3.69	2.46
13	Lots of/Hardly any hunters go on hunting trips when the hunting season has begun/finished for some time.	0.47	0.07	0.47	0	HC	4.5	1.69	4.38	1.77
14	Lots of/Very few cinemas show movies using projectors that are working/broken during the daytime.	0.07	0	0.27	0	LC	4.69	1.92	4.69	1.23
15	Many/Few butchers in the UK slaughter cows in a way that is humane/cruel according to research.	0.4	0	0.27	0.07	HC	3.92	1.77	4.38	2.42
16	A large number of/A small number of burglars break into homes when the owners are gone/home during the evening.	0.07	0.07	0.13	0.07	LC	4.62	2	4.17	1.85

17	Nearly all/Practically no professors teach their regular classes during term-time/holidays as a rule.	0.07	0	0.2	0	LC	4.54	1.67	4.31	1.31
18	Lots of/Hardly any golfers play on the golf course when the weather is sunny/wet over the summer.	0.27	0	0.2	0	LC	4.46	2	4.08	1.67
19	Most/Few tourists visit museums on days when the weather is rainy/sunny on their holiday.	0.07	0	0.33	0.13	LC	4	2.54	4	2.42
20	Almost all/Almost no watchmen like working in neighbourhoods that are safe/dangerous during their shifts.	0.4	0	0.27	0	HC	3.92	2.46	3.58	2.31
21	Nearly all/Rather few shoppers like to buy groceries at stores that seem cheap/expensive when budgets tighten.	0.2	0	0.47	0	HC	4.15	2.08	3.92	1.92
22	Lots of/Hardly any speakers address audiences successfully by talking very clearly/quietly to the room.	0.33	0.07	0.6	0	HC	4.25	1.23	4	1.77
23	Lots of/Very few invaders treat local populations in a way that is cruel/kind as they pillage.	0.27	0	0.4	0	HC	3.62	1.83	4.08	2.62
24	Many/Few gardeners plant their flowers during the spring/winter for best results.	0.6	0	0.73	0	HC	3.77	1.77	4.08	2.67
25	Most/Few architects nowadays design buildings that are efficient/wasteful to save money.	0	0	0	0	LC	4.17	1.92	3.77	2.15
26	Almost all/Almost no negotiators settle disputes faster by behaving politely/aggressively towards both parties.	0.2	0	0.27	0	LC	3.92	2	3.69	2.23
27	Practically all/Practically no witnesses describe robbers as being very hostile/kind in police statements.	0	0	0.13	0	LC	3.46	1.46	4.38	1.5
28	A large number of/A small number of couriers deliver packages and goods on Monday/Sunday in the UK.	0	0	0	0.07	LC	3.62	1.38	4.08	2
29	Nearly all/Rather few brokers sell stocks that are likely to turn a profit/loss for their business.	0.67	0.07	0.33	0.47	HC	4	2.08	3.69	1.85
30	Lots of/Hardly any mechanics fix cars when they are parked in the garage/street during the day.	0.8	0	0.4	0.13	HC	3.69	2.23	3.85	2.08
31	Lots of/Very few joggers like running on a ground that is flat/wet for safety reasons.	0.2	0	0.27	0.07	LC	4.15	1.92	3.83	2
32	Many/Few housewives wash clothes with water that is clean/dirty for optimum results.	0.27	0.07	0.47	0	HC	4.33	1.62	4.54	1.77
33	Almost all/Almost no investors loan money to businesses they think are trustable/hopeless and expect profit.	0.07	0	0.07	0	LC	4.38	1.15	4.5	2.15
34	Practically all/Practically no undertakers arrange funerals with music that is really sombre/cheery for the mourners.	0.13	0	0.13	0	LC	4	1.46	3.92	1.54
35	A large number of/A small number of actors perform plays for which buying tickets is required/free at the booth.	0.07	0	0.13	0	LC	3.92	2.08	3.77	1.92
36	Many/Few adolescents play games on computers that involve cars/cooking after school finishes.	0.07	0	0.13	0	LC	4.15	2	4	2.15
37	Lots of/Hardly any cowboys are found riding their horses in the desert/city during the day.	0.33	0	0.27	0.07	LC	3.42	1.38	4.15	2.23
38	A lot of/A very few barbers create hairstyles that will make their customer look good/bad when they leave.	0.07	0	0.07	0.07	LC	4.46	2.17	4.23	2.15
39	Many/Few bartenders find it easy to hear orders when the bar is quiet/busy during the day.	0.53	0	0.33	0	HC	4.23	2.15	4.54	1.83
40	Almost all/Almost no warehouses store clothes in packaging that is clean/dirty to avoid stains.	0.13	0	0.27	0	LC	3.92	1.92	4	1.54
41	Practically all/Practically no coaches encourage athletes to follow a diet that is healthy/unhealthy for best performance.	0.87	0	0.87	0	HC	4.77	1.62	4.62	1.42
42	A large number of/A small number of plumbers are fixing people's taps when the parts are	0.07	0	0.07	0	LC	4.15	1.92	3.75	2.54

	old/new in the kitchen.									
43	Lots of/Hardly any nurses talk to worried parents in a manner that is calming/rude during difficult times.	0.2	0	0.13	0	LC	4.69	1.85	4.38	1.5
44	Lots of/Very few caterers prepare food that is completely cooked/inedible for dinner parties.	0.27	0	0.27	0.13	LC	4.31	1.38	4.33	1.69
45	Many/Few weathermen encourage people to go to the beach when the temperature rises/drops during the year.	0.53	0	0.27	0.07	HC	4.42	1.62	4.23	1.85
46	Most/Few knights would feel safe wearing armour that was thick/thin during a battle.	0.13	0	0.13	0	LC	3.92	1.54	4.23	1.75
47	Practically all/Practically no freighters are allowed to travel if the cargo is safe/illegal when crossing borders.	0.27	0	0.33	0	LC	4	1.46	4	1.77
48	Nearly all/Rather few policemen say that trying to calm rioting crowds is difficult/easy during a demonstration.	0.47	0	0.53	0.07	HC	4.46	1.23	4.25	1.62
49	Lots of/Hardly any actresses are happy with paparazzi photographs that make them look good/bad in the magazines.	0.27	0	0.2	0	LC	4.42	1.23	4.08	1.69
50	Lots of/Very few adults who are obese wish their weight to decrease/increase for health reasons.	0.4	0	0.67	0	HC	4.46	1.58	4.38	1.46
51	Many/Few teachers would reprimand a student for being silly/clever in the classroom.	0.07	0	0.13	0	LC	4.15	1.77	4.46	2.17
52	Almost all/Almost no scientists conduct research that they think is relevant/pointless to their field.	0.13	0	0.27	0	LC	4.54	1.5	4.38	1.54
53	A large number of/A small number of examiners give their pupils marks that are fair/unfair when grading papers.	0.2	0.13	0.4	0.07	HC	4.23	1.92	4	1.92
54	Almost all/Rather few supervisors disapprove of employees that turn up early/late for a shift.	0.93	0	0.87	0	HC	4.38	1.33	4.31	2.08
55	Lots of/Very few parents would allow their children to play outside if traffic were quiet/busy on their street.	0.27	0	0.47	0	HC	4.31	1.38	4.33	2.31
56	Most/Few businessmen invest all their money in businesses that are reliable/failing to maximise returns.	0.13	0	0.2	0	LC	4.15	1.69	4.15	2.33
57	Practically all/Practically no athletes compete dressed in shoes that are enduring/damaged to avoid injury.	0.07	0	0.13	0	LC	4	1.15	4.23	1.85
58	A large number of/A small number of farms in the UK harvest their crops by machine/hand when they're ready.	0.4	0.07	0.67	0.07	HC	4.46	2.08	4.08	2
59	Nearly all/Rather few countries erect statues for people whose reputation was respectable/poor after their deaths.	0.13	0	0.13	0	LC	3.92	1.46	3.83	2.31
60	A lot of/A very few farmers prepare to milk cattle when the udders are full/empty in the morning.	0.53	0	0.33	0	HC	4.08	1.42	4.38	1.46
61	Most/Few newlyweds would say that their wedding day was a success/disaster after the event.	0.33	0.2	0.27	0.2	HC	4.17	1.31	4	2.08
62	Practically all/Practically no attorneys wear a suit when they are working from their office/home during the week.	0.73	0.13	0.53	0	HC	4.15	2	4	1.83
63	Most/Few engineers become successful by designing devices that work well/badly for the user.	0.47	0	0.53	0.07	HC	4.23	1.15	4.42	2
64	Lots of/Hardly any retirees report that their health is starting to decline/improve as they age.	0.33	0.13	0.33	0.13	HC	4.08	2.77	3.77	2.08
65	Most/Few matadors are chased by a bull when the red flag is moving/still in the air.	0.13	0	0.13	0	LC	3.69	2.69	3.31	2.25
66	Almost all/Almost no Eskimos catch fish in places where it is officially possible/banned according to law.	0.07	0	0.13	0	LC	3.46	2.15	3.17	2.54

67	Practically all/Practically no astronomers go star-gazing when the sky is clear/cloudy late at night.	0.73	0.07	0.73	0	HC	4.25	1.15	3.85	1.62
68	Nearly all/Rather few employees come into the office during the morning/weekend to start work.	0.2	0	0.27	0	LC	4.38	1.62	4.17	2
69	Lots of/Hardly any secretaries invite friends to dine at their home/work at the weekends.	0.13	0	0.13	0.27	LC	3.42	1.85	3.46	2
70	Lots of/Very few surgeons perform operations when their instruments are clean/dirty to avoid infection.	0.67	0.07	0.47	0	HC	4.46	1.75	4.31	1.46
71	Practically all/Practically no sailors prefer to sail when the sea is calm/rough during the day.	0.87	0.07	0.53	0	HC	3.77	1.62	4.54	1.75
72	Nearly all/Rather few songwriters become famous by writing songs that are really catchy/boring to the audience.	0.47	0	0.13	0.07	HC	4	1.67	4	1.77
73	Most/Few therapists are trained to use counselling techniques that are helpful/harmful for the patients.	0.2	0	0.13	0	LC	4.38	1.54	4.46	1.67
74	Almost all/Almost no lifeguards at the beaches swim exceptionally well/badly in open water.	1	0	0.53	0.33	HC	4.62	1.15	4.25	1.31
75	Practically all/Practically no dentists pull out teeth that are clearly rotten/healthy during routine examinations.	0.67	0	0.6	0	HC	4.33	1.15	4.08	1.15
76	A large number of/A small number of salesmen in stores sell designer items that are overpriced/fake to make profit.	0.2	0.07	0.2	0	LC	3.85	1.92	3.46	2.92
77	Nearly all/Rather few aquariums have shark-handling sessions that are safe/dangerous for their visitors.	0.2	0.13	0.47	0	HC	3.77	1.69	4	2
78	Lots of/Very few tenants think that the rent they pay is too high/low for the flat.	0.73	0	0.87	0.13	HC	4.08	1.75	3.77	1.92
79	Most/Few parents let their children play outside when it's sunny/dark in the streets.	0.47	0	0.8	0	HC	4.58	1.38	4.23	1.54
80	Almost all/Almost no people go sales shopping to find something cheap/expensive on Boxing Day.	0.4	0	0.6	0	HC	4.38	1.62	4.38	2
81	Lots of/Very few students find theoretical maths problems very difficult/easy to work out.	0.53	0	0.6	0.13	HC	4.08	1.92	3.62	1.69
82	Lots of/Rather few accidents in traffic result from young men driving too fast/slowly on the road.	0.67	0.07	0.33	0.13	HC	4.23	1.69	4	2
83	Very many/Very few people think that giving roses on Valentine's day is typical/original as a surprise.	0.07	0	0.13	0	LC	3.92	1.38	4.38	1.54
84	Lots of/Very few people enjoy walking somewhere that is peaceful/crowded to help relax.	0.13	0	0.13	0	LC	4.08	2.08	3.62	2
85	Almost all/Very few people drive their car when they are completely sober/drunken to avoid accidents.	0.73	0	0.53	0	HC	4.77	1.54	4.42	1.08
86	Lots of/Very few people light a wood fire when the weather is cold/hot throughout the year.	0.8	0	0.2	0	HC	3.83	1.46	4	2.31
87	Most/Few students of quantum physics are known for being diligent/lazy by their professors.	0	0	0	0	LC	3.92	1.77	3.83	2.85
88	Lots of/Hardly any people are happy to eat chicken that is cooked/raw for their dinner.	0.33	0	0.4	0	HC	4.42	1.23	4.54	1.85
89	Almost all/Very few plays during the Edinburgh festival are pretty fascinating/rubbish for their audiences.	0.07	0	0.07	0	LC	3.38	2.33	3.69	2
90	A large majority of/A small minority of children in the UK prefer food that is unhealthy/healthy at each mealtime.	0.47	0	0.53	0.07	HC	4.08	2.38	3.92	2.08
91	Most/Few countries in the third world have governments that are unreliable/stable by Western standards.	0.07	0	0.2	0	LC	4.17	1.46	3.92	1.85

92	Nearly all/Rather few Picassos on the black market for art are fake/real according to experts.	0.67	0	0.33	0.07	HC	3.77	1.92	4.23	2
93	Most/Few people would be happy to overnight become wealthy/poor by today's standards.	0.13	0	0.13	0.07	LC	4.69	1.85	4.54	1.42
94	Most/Few people would be happy if their architect and builder were reliable/frauds working without accreditation.	0.07	0	0.07	0	LC	4.46	1.42	4.38	1.38
95	Mostly all/Hardly any people consider Neanderthals to be truly primitive/civilised by modern standards.	0	0	0	0	LC	3.92	2.85	4.17	2.46
96	All/No parents wish their own children to be polite/rude to other people.	0	0	0.07	0	LC	4.17	1.15	4.08	1.54
97	Almost all/Almost no people believe that Harry Potter is imaginary/real except young children.	0.13	0.07	0.93	0	HC	4.08	1.42	4.38	2.23
98	Lots of/Few people prefer to shop when the shops are empty/busy in the evenings.	0.33	0	0.4	0	HC	4.33	1.92	3.92	2.31
99	Lots of/Few people injure themselves by diving into water that is shallow/deep at the beach.	0.87	0	0.8	0	HC	3.92	2.08	3.54	1.54
100	Very many/Very few children love food that is really unhealthy/healthy to parents' frustration.	0.2	0	0.6	0	HC	4	2.54	3.69	2
101	Almost all/Almost no collectors of antiques look for something unique/common when attending auctions.	0.33	0	0.27	0	HC	4.5	1.54	4.08	1.38
102	Many/Few girls like to date men who are relatively handsome/ugly compared to most.	0.13	0	0.13	0	LC	4.46	1.92	4.23	1.58
103	Most/Few people feel safe at home when the door is locked/open late at night.	0.93	0	0.47	0	HC	4.15	2.33	3.92	1.62
104	Lots of/Hardly any industries who dispose of waste illegally are caught eventually/immediately by the authorities.	0.13	0.07	0.2	0	LC	3.85	2.38	4	2.38
105	Very many/Almost no people think that unemployment will begin to increase/decrease in coming years.	0.13	0.2	0.27	0.07	LC	3.54	2.77	3.23	2.17
106	Very many/Very few people find the idea of health care for all to be sensible/bad for today's society.	0	0	0.13	0	LC	4	1.62	3.62	1.92
107	Nearly all/Rather few aeroplanes take off and land on runways/water in the UK.	0.33	0	0.67	0	HC	4.69	2	4.38	1.31
108	Many/Few politicians answer sensitive questions in a manner that is misleading/direct according to voters.	0.13	0	0.2	0	LC	3.62	2.08	3.62	1.85
109	Almost all/Almost no adults find that teenage boys are very moody/mature especially at home.	0.07	0	0.13	0	LC	4	1.69	3.58	1.92
110	Most/Few gamblers who play roulette in casinos will eventually lose/win given the odds.	0.67	0.07	0.73	0.13	HC	4.42	2.38	3.92	2.15
111	Most/Few people admire those performers who are very good/bad at their craft.	0.13	0	0.2	0	LC	4.23	1.62	4.08	1.83
112	Lots of/Very few people with mental health problems find getting a job difficult/easy because of prejudice.	0.53	0.13	0.73	0.13	HC	4.62	1.62	4.42	2.23
113	The majority of/A minority of Brits believe that men and women should be treated equally/differently in today's society.	0.87	0	0.27	0.47	HC	4.38	1.92	4.08	1.54
114	Nearly all/Hardly any parents find the first weeks of child-rearing difficult/easy to cope with.	0.53	0.07	0.8	0	HC	4.08	1.42	4.23	1.46
115	Practically all/Practically no people find badly burnt food to be disgusting/tasty to swallow down.	0.27	0	0.4	0	HC	3.67	1.62	4.38	1.77
116	Most/Few people find tying their shoelaces to be easy/difficult beyond age seven.	0.47	0.07	0.67	0.2	HC	4.15	2.25	4.08	1.38
117	Most/Few students of English literature find poetry to be meaningful/meaningless as a pastime.	0.07	0	0	0	LC	4.08	2.23	3.62	2

118	Nearly all/Rather few people who star in staged reality television shows are stupid/smart according to viewers.	0.13	0	0.2	0	LC	3.54	2	3.83	2.31
119	Most/Few toddlers who have just woken up are grumpy/happy towards their parents.	0.2	0.07	0.27	0	LC	3.62	2.42	3.77	2
120	Practically all/Practically no Rolexes that are sold on the street are fake/real in these times.	0.73	0	0.53	0	HC	4.46	1.69	4.54	1.75
121	Almost all/Almost no candidates for important job interviews feel uncomfortable/comfortable in that situation.	0	0	0	0	LC	4.15	1.77	4.38	2.33
122	Practically all/Practically no people who bungee-jump for the first time feel scared/relaxed right before jumping.	0.27	0	0.27	0	LC	4	1.92	4	1.31
123	Most/Few nurses who care for the elderly are female/male in British hospitals.	0	0	0.07	0	LC	3.69	2	3.54	1.85
124	Most/Few people who get a filling prefer to be asleep/awake during the procedure.	0.2	0	0.2	0.2	LC	3.46	2.69	3.46	2.58